Office of Naval Research
Contract N00014-88-K-0472
R T Code 413p008

Technical Report No. 5

Molecular Modeling Studies of Cholic Acid Host: Shape And Dimensions. II.

by

Suzanne M. Evans* and Carol A. Venanzi

New Jersey Institute of Technology
Chemistry Division
Newark, NJ 07102

May 21, 1990

*: The BOC Technical Center, Murray Hill, NJ

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION New Jersey Institute of Technology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Chemistry Division University Heights Newark, New Jersey 07102	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 2217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Molecular Modeling Studies of Cholic Acid Host: Shape and Dimensions. II.

12. PERSONAL AUTHOR(S)
Suzanne M. Evans and Carol A. Venanzi

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
	FROM _____ TO _____	May 21, 1990	18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Molecular Structure, Bile, Models, Shape, Xylenes, Aromatic Compounds

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Earlier studies Technical Report No. 3 on the three-dimensional "clam" or cup shapes formed by two cholic acid molecules linked together by a substituted aromatic spacer had shown that the resulting cavities were of similar shape but different width, due to the flexible nature of this host system, with its sensitivity to minor bond rotations. In the L-arabinose binding protein, the binding pocket shape is cup-like, approximately 12-12.3 Å long and 7.3 Å wide (taken from crystallographic coordinates, Fig. 1). In order to mimic the sugar binding proteins, wherein the sugar hydroxyl groups are hydrogen bonded to the ionized amino acids of the protein through the heteroatoms, (with or without an intervening water network, Fig. 2) it was necessary to investigate the dimensions of the cholic acid host cavities more closely.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Harold E. Guard	22b. TELEPHONE (Include Area Code) (202) 696-4402 22c. OFFICE SYMBOL ONR

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

19. ABSTRACT -continued-

The present study extends th- earlier report by defining a "standard" for the cholic acid host structure, one that uses two dihedral angles, each CH_2 to the aromatic ring, to define the origin direction of the two arms (or halves) of the clam shapes. The definition is applied to cholic acid host systems containing a m-xylylene and a p-xylylene aromatic spacer. Subsequent manipulation of two critical dihedral angles in the flexible portion of the hosts creates cavities which are then evaluated geometrically by measuring the interatomic distances between key heteroatoms (hydroxyls) in each arm. The 90 and 120 degree clam shapes of the m-xylylene cholic acid host are the closest approximations to the L-arabinose binding protein, although these shapes are not perfect "clams" since they have the top arm offset relative to the bottom arm. In the p-xylylene cholic acid hosts, two 90 degree structures seem to create good clam shapes in which to dock a monosaccharide guest. Geometry optimization with energy minimization of these two hosts results in very little change in the cavity shapes but does relieve several bad steric interactions. The size of the p-xylylene cavities are approximately 12 Å long an 9-11 Å wide.

The Cholic Acid Host System Shape

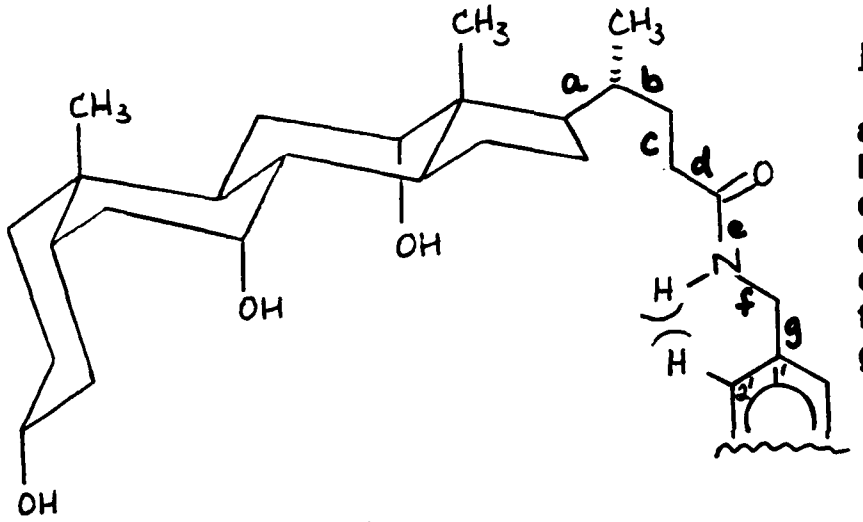
Abstract

Earlier studies on the three-dimensional "clam" or cup shapes formed by two cholic acid molecules linked together by a substituted aromatic spacer had shown that the resulting cavities were of similar shape but different width, due to the flexibile nature of this host system, with its sensitivity to minor bond rotations (see report entitled The Cholic Acid Host System Shape of 2/25/90). In the L-arabinose binding protein, the binding pocket shape is cup-like, approximately 12-12.3 Å long and 7.3 Å wide (taken from crystallographic coordinates, Fig. 1). In order to mimic the sugar binding proteins, wherein the sugar hydroxyl groups are hydrogen bonded to the ionized amino acids of the protein through the heteroatoms (with or without an intervening water network, Fig.2), it was necessary to investigate the dimensions of the cholic acid host cavities more closely.

The present study extends the earlier report by defining a **"standard"** for the cholic acid host structure, one that uses two dihedral angles, each CH_2 to the aromatic ring, to define the origin direction of the two arms (or halves) of the clam shapes. The definition is applied to cholic acid host systems containing a m-xylylene and a p-xylylene aromatic spacer. Subsequent manipulation of two critical dihedral angles in the flexible portion of the hosts creates cavities which are then evaluated geometrically by measuring the interatomic distances between key heteroatoms (hydroxyls) in each arm. The 90 and 120 degree clam shapes of the m-xylylene cholic acid host are the closest approximations to the L-arabinose binding protein, although these shapes are not perfect "clams" since they have the top arm offset relative to the bottom arm. In the p-xylylene cholic acid hosts, two 90 degree structures seem to create good clam shapes in which to dock a monosaccharide guest. Geometry optimization with energy minimization of these two hosts results in very little change in the cavity shapes but does relieve several bad steric interactions. The size of the p-xylylene cavities are approximately 12 Å long and 9-11 Å wide.

Meta-xylylene Aromatic Spacer

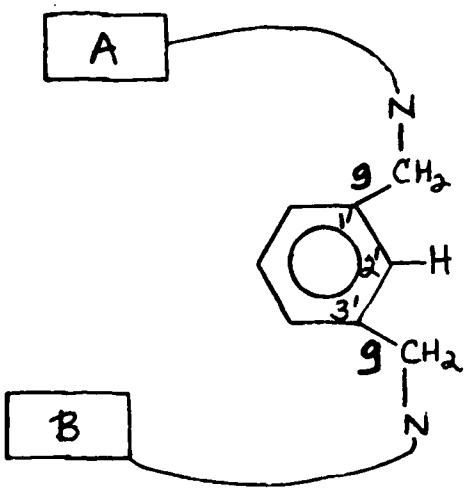
The torsional angle values for bonds labeled **a**, **b**, **c**, **d**, and **e** were obtained from crystallographic coordinates (refcode = BUGJES), while **f** and **g** were set to create the reference conformation shown below. While **a**, **c**, **d**, and **e** remained fixed in both halves of the cholic acid host containing the meta-xylylene spacer, dihedral angles **b** (C_{17}--C_{20}--C_{22}--C_{23}) and **g** (benzylic C--$C_{1'}$--$C_{2'}$--H_{ar}) were varied. Dihedral angle **b** has the most direct effect on the cholic acid arm vector (directionality), whereas dihedral angle **g** alleviates the bad steric interactions, found in the reference torsional angle of 0 degrees, between the hydrogen of the N-H in the amide and the ortho-proton $C_{2'}$-H of the benzene ring.



Reference

a = +173.6
b = +60.6
c = +179.3
d = -168.7
e = +172.6
f = +180.0
g = 0

The **"standard"** for the cholic acid host is defined as shown: The molecule starts at the left, proceeding along arm **A** through the aromatic spacer toward arm **B**, always with a negative **g** dihedral angle in arm **A**, using $C_{1'}$ and $C_{2'}$. Dihedral angle **g** to arm **B** is defined using $C_{3'}$ and $C_{2'}$.



Arm A-
g defined as
N-C - $C_{1'}$ - $C_{2'}$ NEGATIVE

Arm B-
g defined as
N-C-$C_{3'}$-$C_{2'}$

The following table of dihedral angle combinations thus results from the manipulation of dihedral angle **g**.

Name	Dihedral Angles			
	Bond 20-22 (**b**)		Bond **g** to Aromatic Ring	
	Arm A	Arm B	Arm A	Arm B
_60 Clam	+60.6	+60.6	-60	+60
_90 Clam	+60.6	+60.6	-90	+90
_120 Clam	+60.6	+60.6	-120	+120
_150 Clam	+60.6	+60.6	-150	+150

Each of the four cholic acid hosts above have the top arm **A** offset relative to the bottom arm **B**, thereby creating the effect of hinge movement. Although the general shape is that of a "C" or clam, the cavities are of different dimensions (Fig. 3). The interatomic distances (in Å) between the cholic acid hydroxyls in each arm reflect the cavity size, which differ mainly in width.

Name — Å Distance

_60 Clam

Arm **A** // **B**	O3	O7	O12
O3	7.31	7.73	5.18
O7	10.12	10.09	5.81
O12	7.31	7.75	3.97

_90 Clam

A // **B**	O3	O7	O12
O3	8.66	10.87	7.51
O7	11.98	12.54	8.23
O12	10.64	11.48	7.84

_120 Clam

A // **B**	O3	O7	O12
O3	13.77	15.75	12.11
O7	15.72	16.21	12.08
O12	14.85	15.28	11.86

_150 Clam

A // **B**	O3	O7	O12
O3	18.44	19.67	16.34
O7	19.38	19.38	15.77
O12	17.97	17.77	14.91

Comparing the above synthetic cavities to the dimensions of L-arabinose binding protein, 12 Å long and 7Å wide, the _60 Clam is clearly too small in the interior. In addition, this clam shape has its ends (A-rings of the two steroidal arms) overlapping. The _150 Clam shape clearly has the largest cavity, although this may indeed be too large for a monosaccharide to maximize its hydrogen bonding interactions with both steroidal arms. The _120 and _150 Clam structures, however, seem to be reasonable

cavities in which to dock a sugar guest , with the _90 Clam most closely resembling the cavity in L-arabinose binding protein.

Manipulation of dihedral angles **b** (C_{20}-C_{22}) and **g** were done in several structures to evaluate the effect on the cavity accompanying changes in the two most crucial bonds of the flexible chain of each cholic acid arm. The following table shows the single modifications which were made.

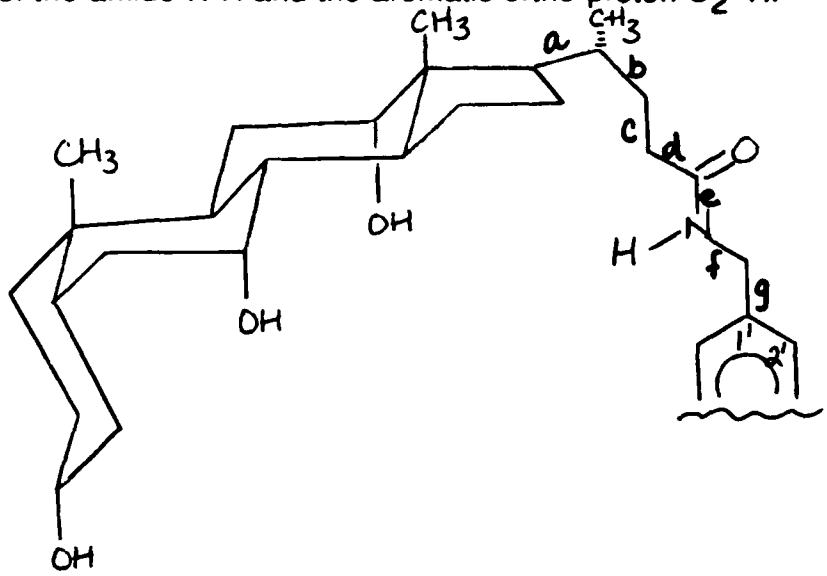
	Dihedral Angles			
Name	Bond 20-22 (**b**)		Bond **g** to Aromatic Ring	
	Arm A	Arm B	Arm A	Arm B
_90b	-60.6	+60.6	-90	+90
_90bE	-60.6	+60.6	+90	+90
_90E	+60.6	+60.6	+90	+90
_120E	+60.6	+60.6	-120	-120
_150MI*	-60.6	-60.6	+150	-150

* the total mirror image, created by reflection of X-coordinates through YZ plane

The _90b has a negative dihedral angle **b** and does not create a clam shape since the hydroxyls on arm **A** are directed outward instead of inward (Fig. 4). Structure _90bE, the extended form of _90b, and _90E, _120E, the extended forms of _90 Clam and _120 Clam respectively, illustrate that an open conformation (as opposed to a folded conformation) can be created with a single dihedral angle modification, namely **g** in arm **A** or **B** (Fig. 5, _90Clam and _90E). The total mirror image of one conformation, the _150 Clam, shows the same size cavity but opposite orientation in space (Fig. 6). None of these alternate conformations wll be considered as candidates for docking the sugar guests.

Para-xylylene Aromatic Spacer

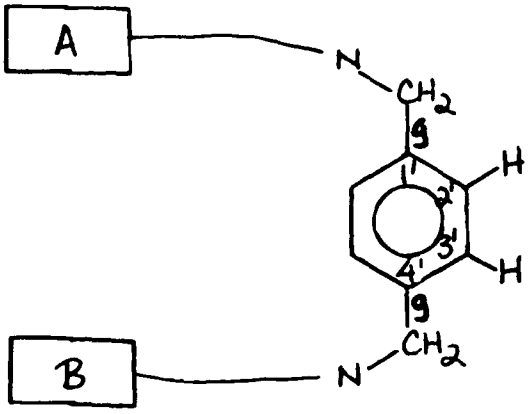
As with the m-xylylene spacer, the torsional angles for bonds labeled **a**, **b**, **c**, **d**, and **e** were obtained from crystallographic coordinates (refcode = CUBSUN for **a**, modification of others to mimic BUGJES), while **f** and **g** were set to create the reference conformation shown below. Dihedral angles **a**, **c**, **d**, **e**, and **f** remained fixed in both halves of the cholic acid host, while **b** (C_{17}--C_{20}--C_{22}--C_{23}) and **g** (benzylic C--$C_{1'}$--$C_{2'}$--H_{ar}) were varied. Again, dihedral angle **b** has the most direct effect on the cholic acid arm vector (directionality), whereas dihedral angle **g** alleviates the bad steric interactions, found in the reference torsional angle of 0 degrees, between the hydrogen of the amide N-H and the aromatic ortho-proton $C_{2'}$-H.



Reference

a = -170.3
b = +60.6
c = +179.3
d = +180.0
e = +180.0
f = +180.0
g = 0

The **"standard"** for this cholic acid host is defined as shown: The molecule starts at the left and proceeds along arm **A** through the aromatic spacer toward arm **B**, always with a negative **g** dihedral angle in arm **A**, using C1' and C2'. This dihedral angle is comparable to that used in the "standard" for the m-xylylene spacer. Dihedral angle **g** in arm **B** is measured using $C_{4'}$ and $C_{3'}$.



Arm A ~
g defined as
N-C-$C_{1'}$-$C_{2'}$ NEGATIVE

Arm B ~
g defined as
N-C-$C_{4'}$-$C_{3'}$

The p-xylylene aromatic spacer provides a more symmetrical clam shape due to the para substitution of the steroidal arms onto the benzene ring. In addition, the N-H to C_2-H interaction is not as severe. Therefore, since the considerations of geometry were different here, the studies with this host proceeded differently than with the m-xylylene fragment. Two combinations of dihedral angles **b** and **g** giving folded "C" shapes were studied, along with two extended "S" shapes. The starting structures in this series were optimized energetically (full geometry, no electrostatics) using the MAXIMIN2 force field found in the SYBYL Molecular Modeling Software (Appendix A). The resulting "C" shapes widened in the middle while moving closer together at the ends, the A rings of the steroidal arms. The "S" shapes tightened in their shape by curving more toward the spacer. The following table shows both the beginning and final energies and dihedral angles for each of the p-xylylene cholic acid hosts.

Name	Energy	Dihedral Angles for Arm A, Arm B	
		Bond 20-22 (b)	Bond g
_PX	+817	+60.6, +60.6	-90, +90
_PX_MAX[a]	+65.1	+62.6, +62.7	-82.6, +97
_PX_0	+1028	+60.6, 0	-90, +90
_PX_0_MAX	+71.06*	+59.9, -35.4	-73.3, +105
_PX_OPEN[b]	+817	+60.6, +60.6	-90, -90
_PX_OPEN_MAX	+67.16	+63.5, +63.7	-85.5, -85.7
_PX_0E[c]	+1029	+60.6, 0	-90, -90
_PX_0E_MAX	+74.54**	+56.6, -38.0	-79.5, -71.8

* +0.04 away from convergence at +71.023

** +0.04 away from convergence at +74.40

[a] _MAX = MAXIMIN2 result, all _MAX results were fit to _PX

[b] the "open" or extended form of _PX

c the extended form of _PX_0

The components of the total energy are given in Appendix B. The initial difference between the two 90 degree clam shapes is shown in Fig. 7. Initial and final structures _PX_0 and _PX_0_MAX are shown in Fig. 8, illustrating the typical change in the "C" clam shape. The structure of _PX_MAX is shown in Fig. 9. Finally, Fig. 10 shows how minimization alters the "S" shape of an extended p-xylylene cholic acid host.

The atom pair distances indicative of the cavities created in the two "C" shapes (both before and after minimization) were:

Name	Å Distance	Arm A // B	O3	O7	O12	Cavity Width (Å)
		O3	10.00	12.72	9.34	
_PX		O7	12.67	14.72	11.29	8.81
		O12	9.26	11.27	8.81	
		A // B	**O3**	**O7**	**O12**	
		O3	9.98	12.68	9.12	
_PX_MAX		O7	12.63	14.86	11.41	9.08
		O12	9.04	11.40	9.08	
		A // B	**O3**	**O7**	**O12**	
		O3	3.01	6.54	8.50	
_PX_0		O7	7.48	8.28	10.57	6.47, 10.57*
		O12	7.78	6.47	9.63	
		A // B	**O3**	**O7**	**O12**	
		O3	2.90	6.56	8.32	
_PX_0_MAX		O7	7.31	9.20	10.85	6.66, 10.85
		O12	7.78	6.47	9.63	

*cavity is larger in the middle than at the ends

The two 90 degree clam shape cavities formed from the p-xylylene aromatic spacer, with widths of 9-11 Å and lengths of 12 Å, are reasonable starting points for incorporating the monosaccharide moiety. The energy minimizations did not significantly alter the cavity sizes but optimized the overall geometry (Fig. 8). The extended forms of these hosts are created by rotation about a singe bond, **g** in arm **B**, from the flexible chain to the aromatic ring. If the molecule exists in an extended form, for example conformation _PX_OPEN_MAX, the distance from the end of the steroid arm **B** to the plane of the xylylyl fragment is 12.03 Å (Fig. 11). Therefore, to create a clam or "C" shape, this arm would have to swing around on the hinge to cover almost 25Å!

To complete this phase of the cholic acid host system geometry evaluation, energy minimizations on the m-xylylene structures must be done and the accompanying change in cavity size and/or shape must be studied. Docking a monosaccharide guest, n-pentyl beta-D-glucoside, has begun in the p-xylylene cholic acid hosts by first visually docking the sugar and then optimizing the whole complex using MAXIMIN2. These studies are preliminary since only the steric (no electrostatic) component to the energy has been included in the evaluation of the host guest complexes.

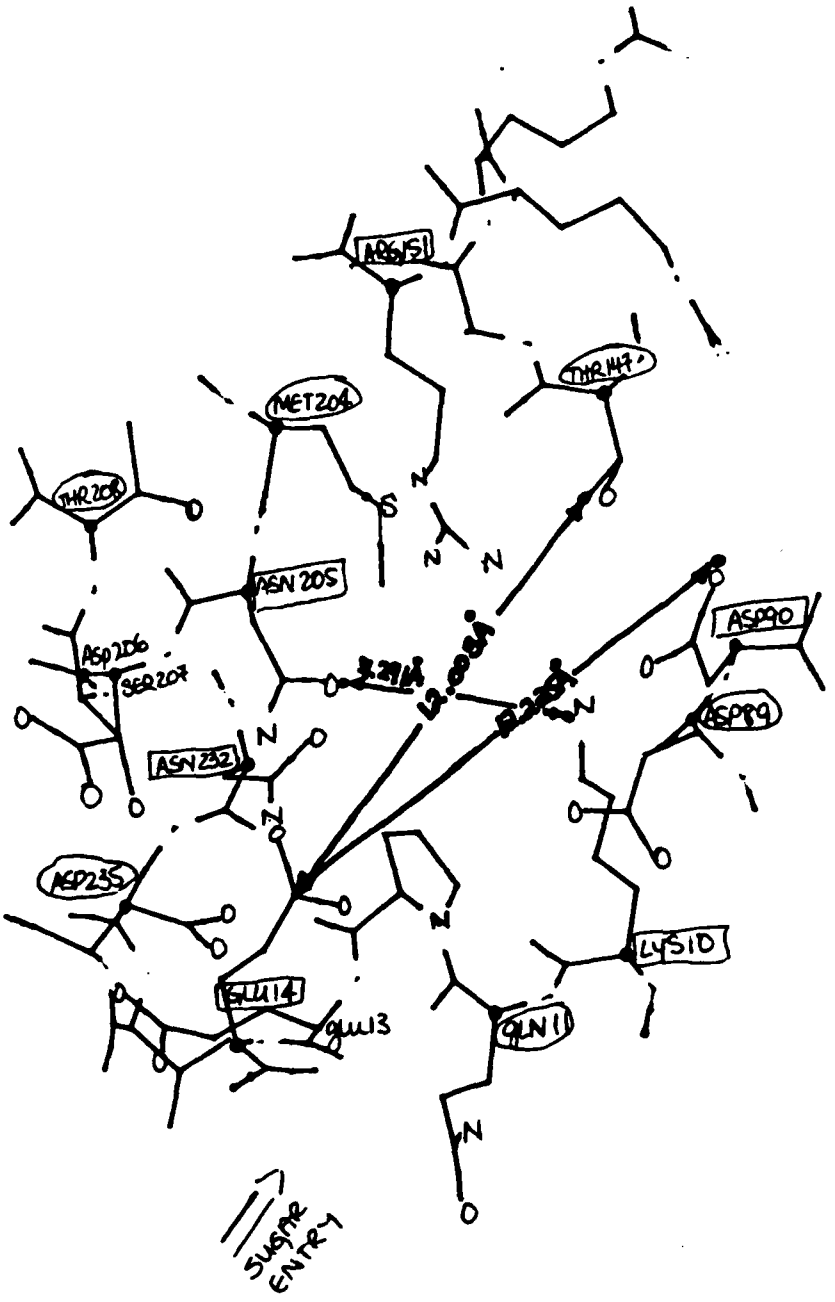


FIGURE 1.

L_ABP_12345 BY_ATOM_COLOR NOV 08 1989

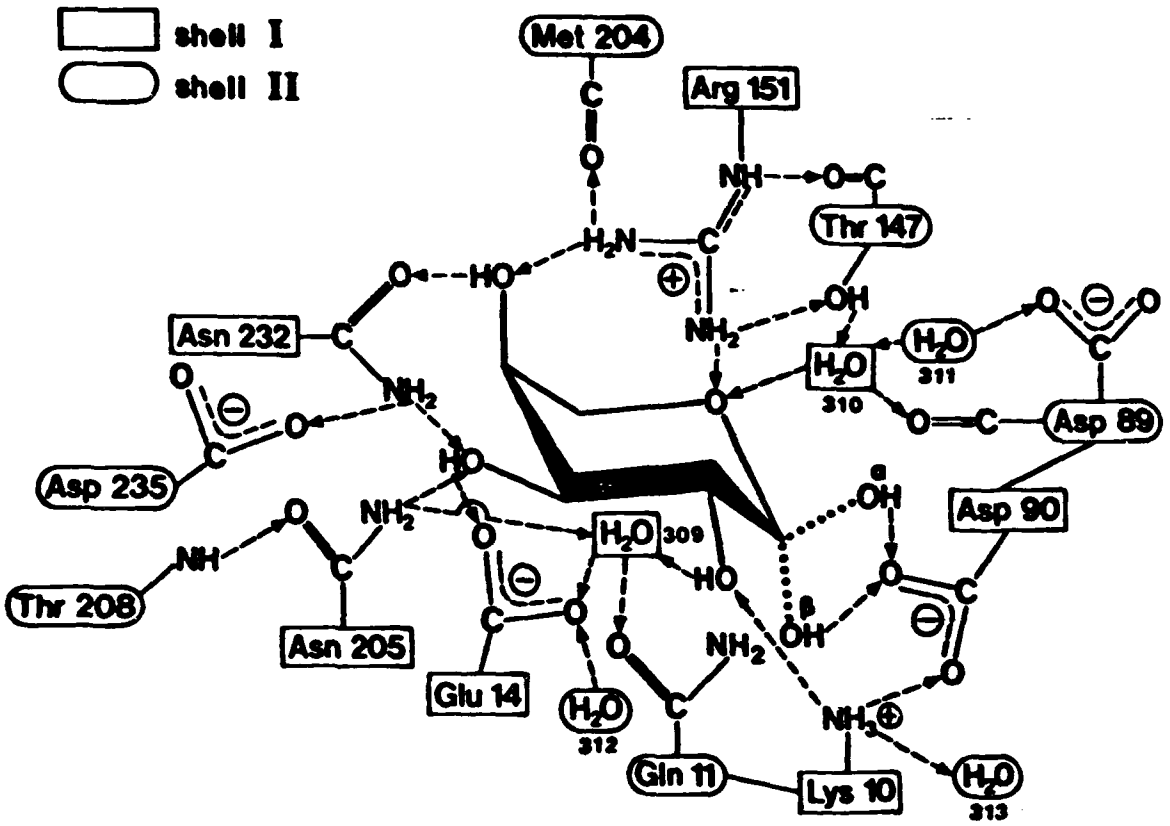


Figure 2. Schematic diagram of the intricate networks of hydrogen bonds formed in the complex of the L-arabinose-binding protein with the L-arabinose substrate. Shell I represents the essential residues hydrogen-bonded to the sugars and to adjacent second shell (shell II) of residues. Note especially that Arg-151 is not involved in salt linkages but is the source of five donor groups for five hydrogen bonds. (Adapted from Quiocho and Vyas 1984.)

FIGURE 2.

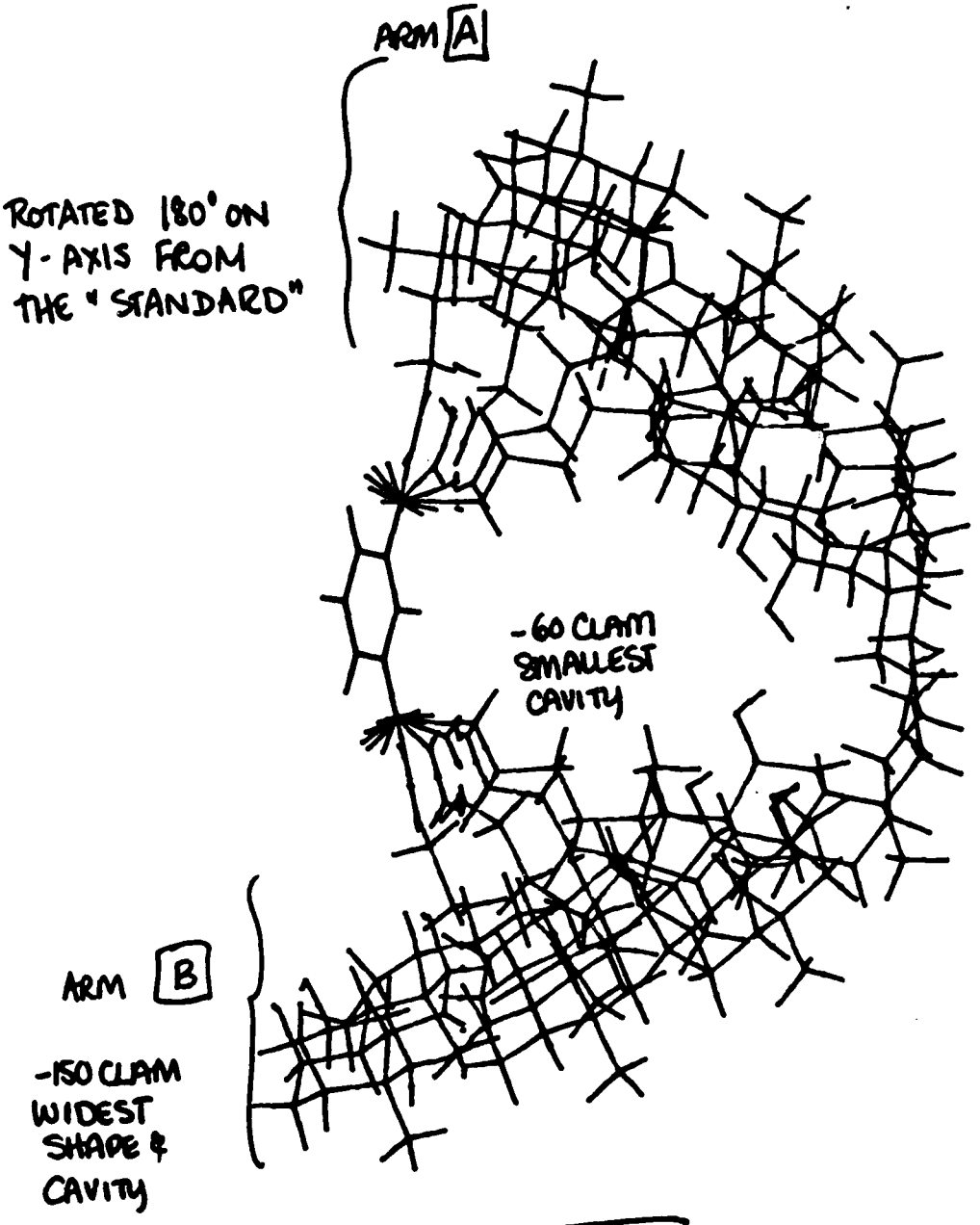


FIGURE 3

ALL BUGJES_MX2_* FITTED

APR 06 1990

TOTAL OF 5 STRUCTURES

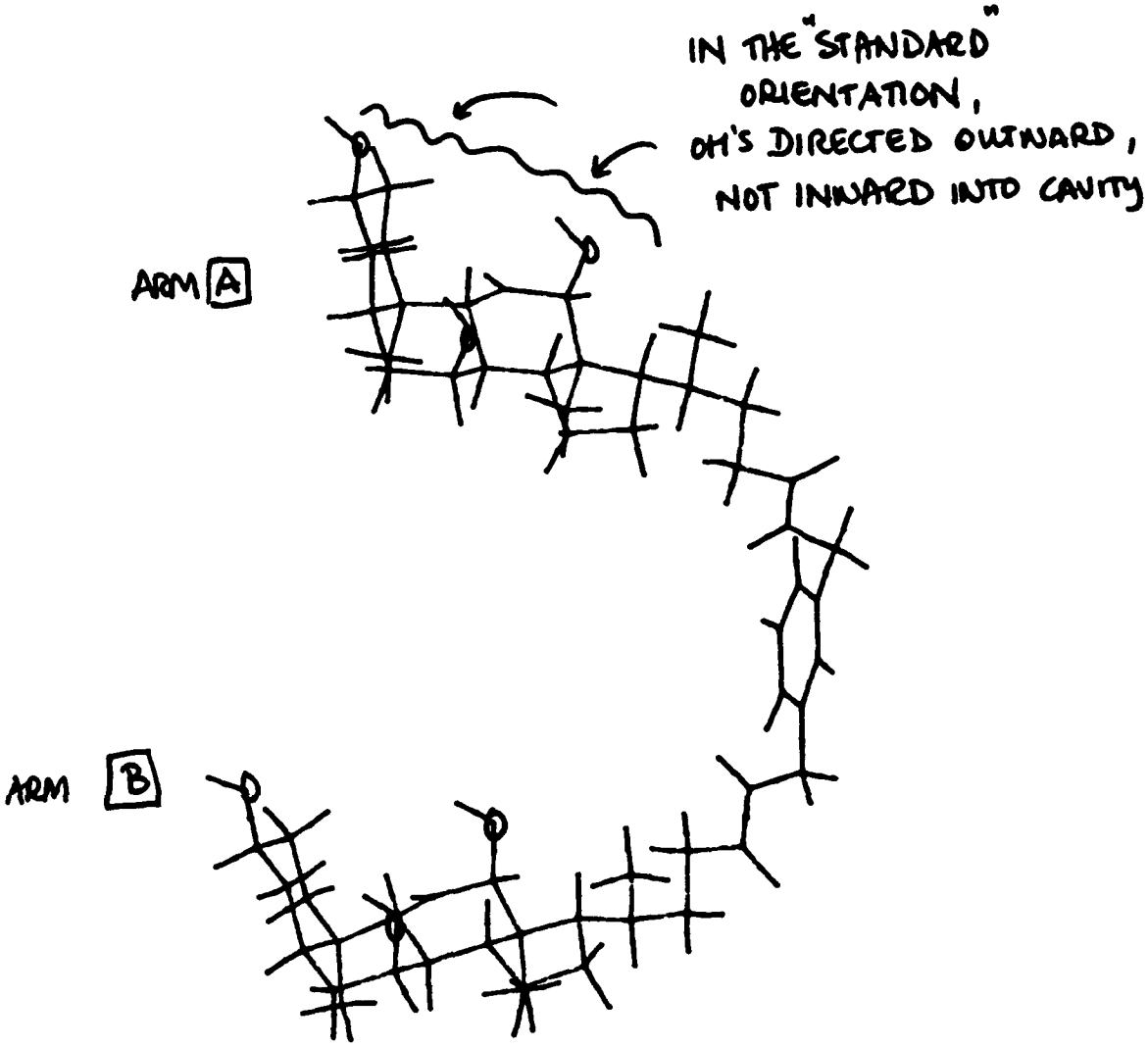


FIGURE 4.

BUGJES_MX2_90-b

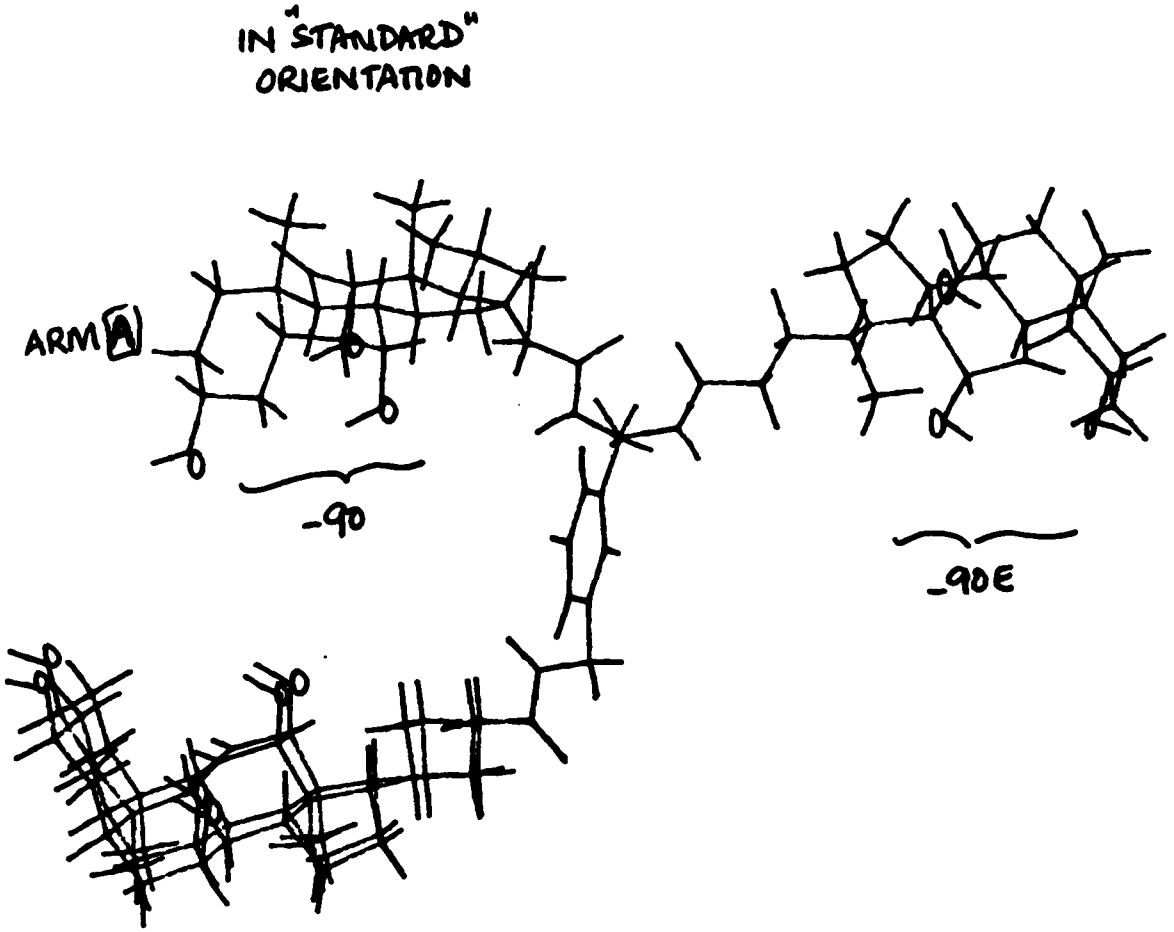


FIGURE 5.

BUGJES_MX2_90E (EAST) ON _90

APR [illegible]

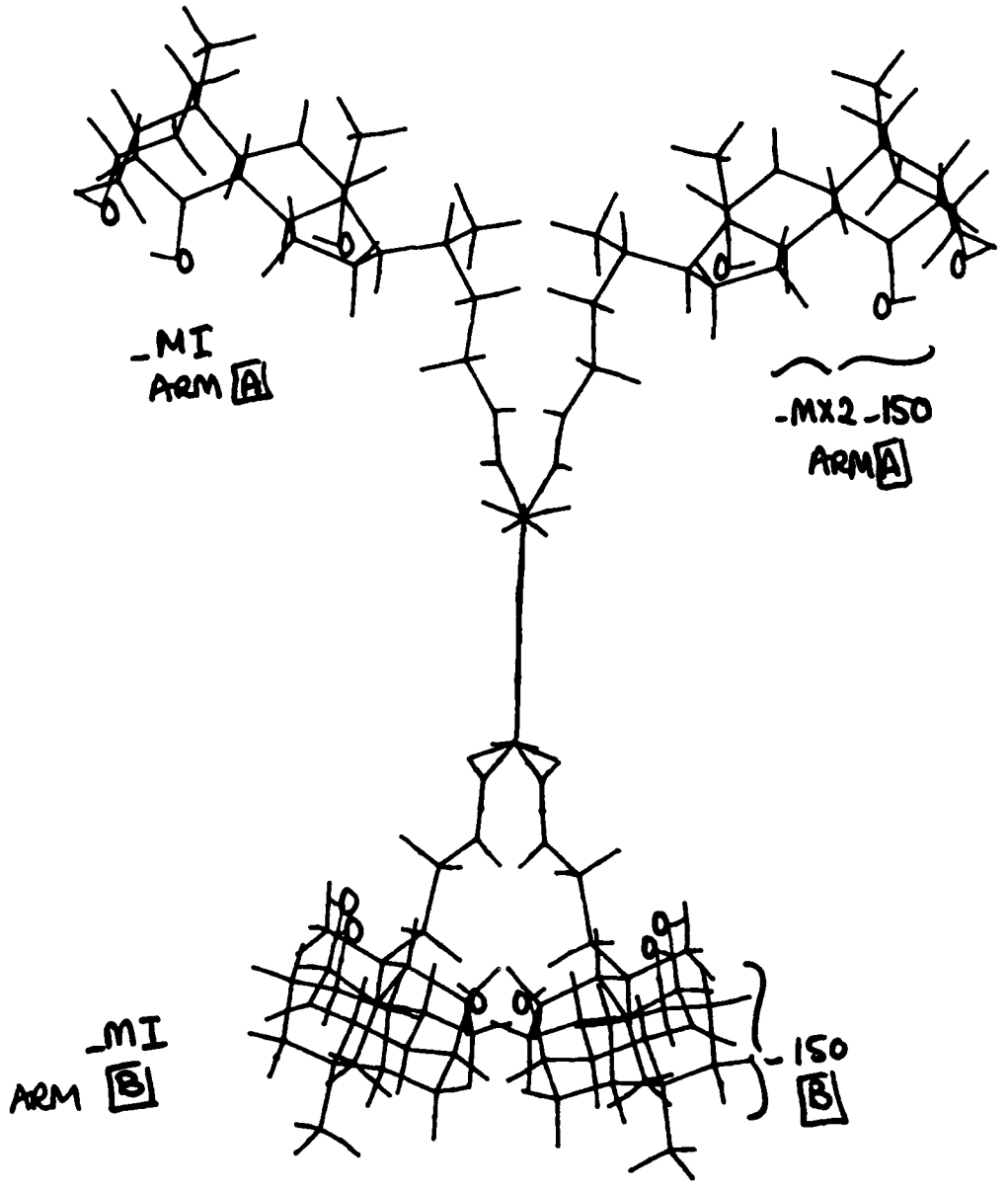


FIGURE 6.

BUGJES_MX2_150 EAST ON _MX2_150_MI WEST

MAY 17 1990

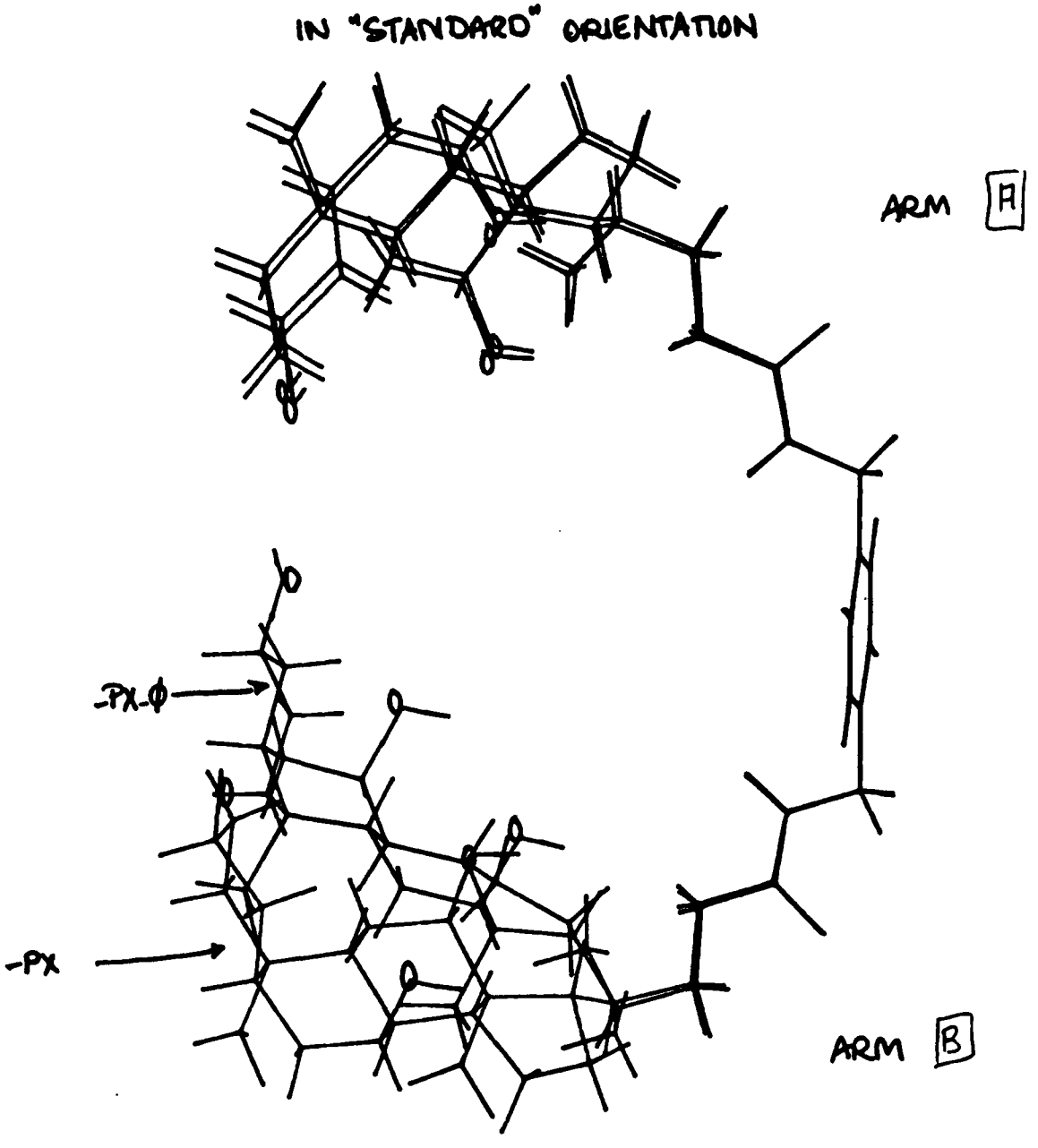


FIGURE 7.

CUBSUN_PX & CUBSUN_PX_0 (RED) ARM A →ARM B

APR [illegible]0

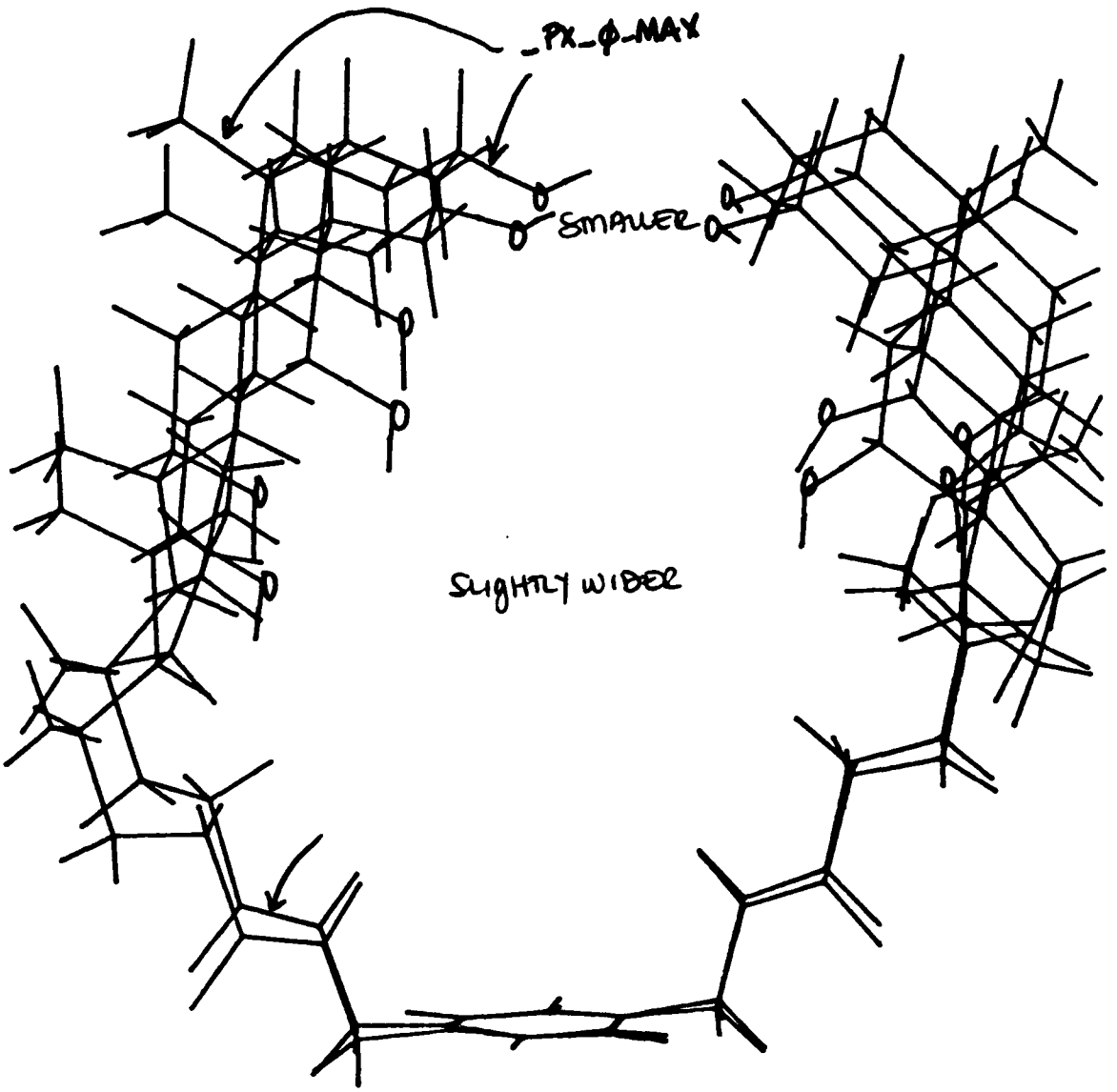


FIGURE 8.

CUBSUN_PX_φ_MAX (RED) ON _PX_φ

APR

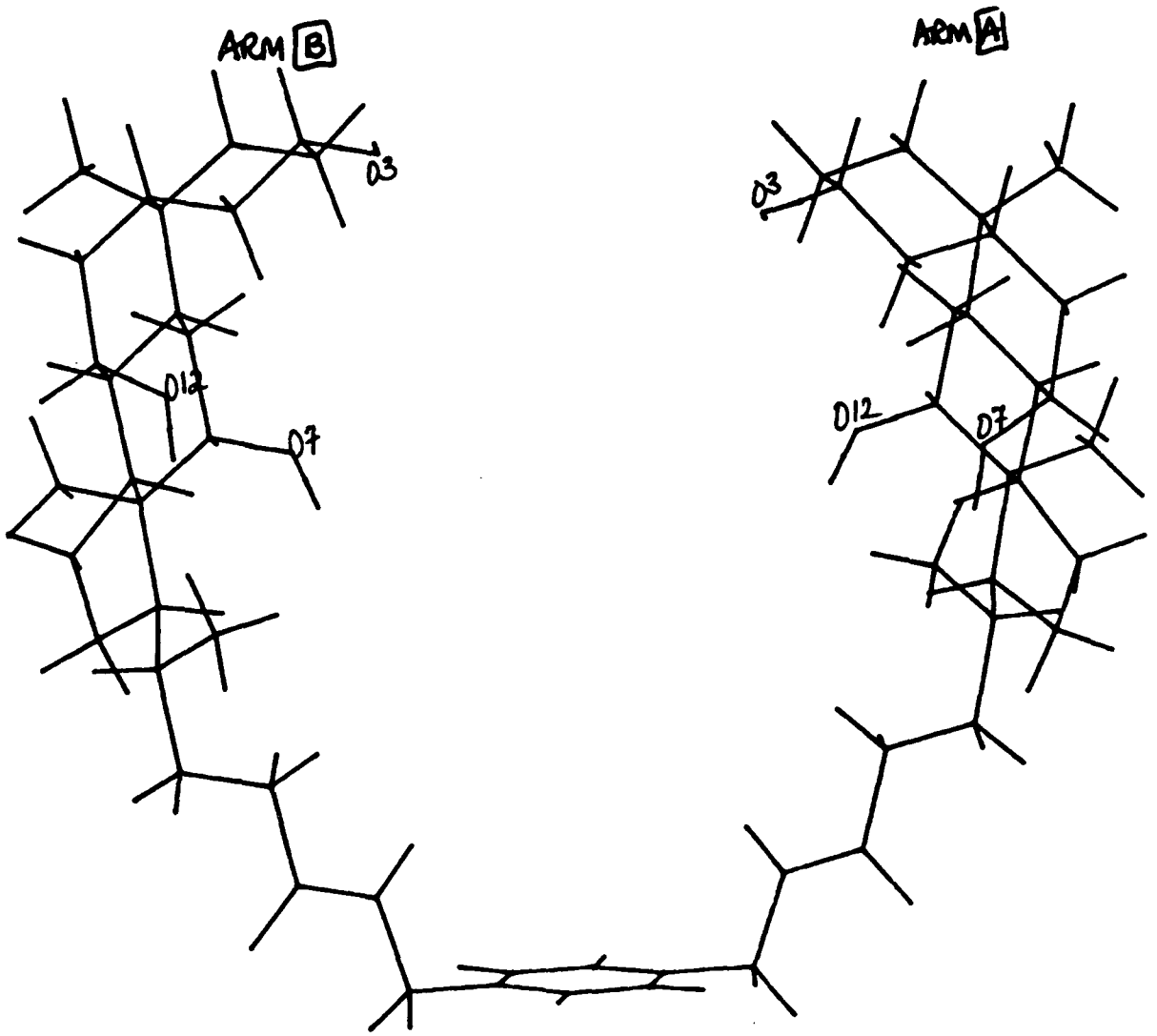


FIGURE 9.

CUBSUN_PX_MAX

APR

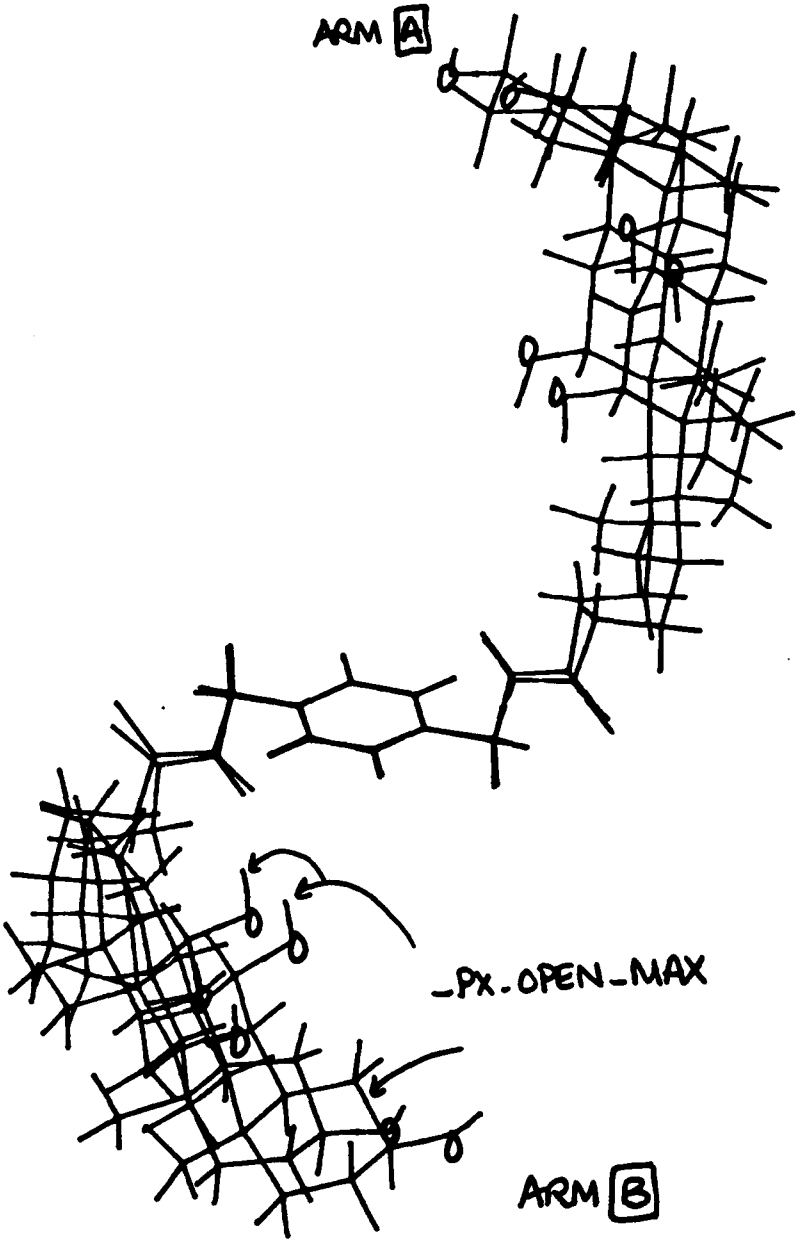


FIGURE 10.

CUBSUN_PX_OPEN_MAX ON GREEN INPUT

APR [illegible] 1990

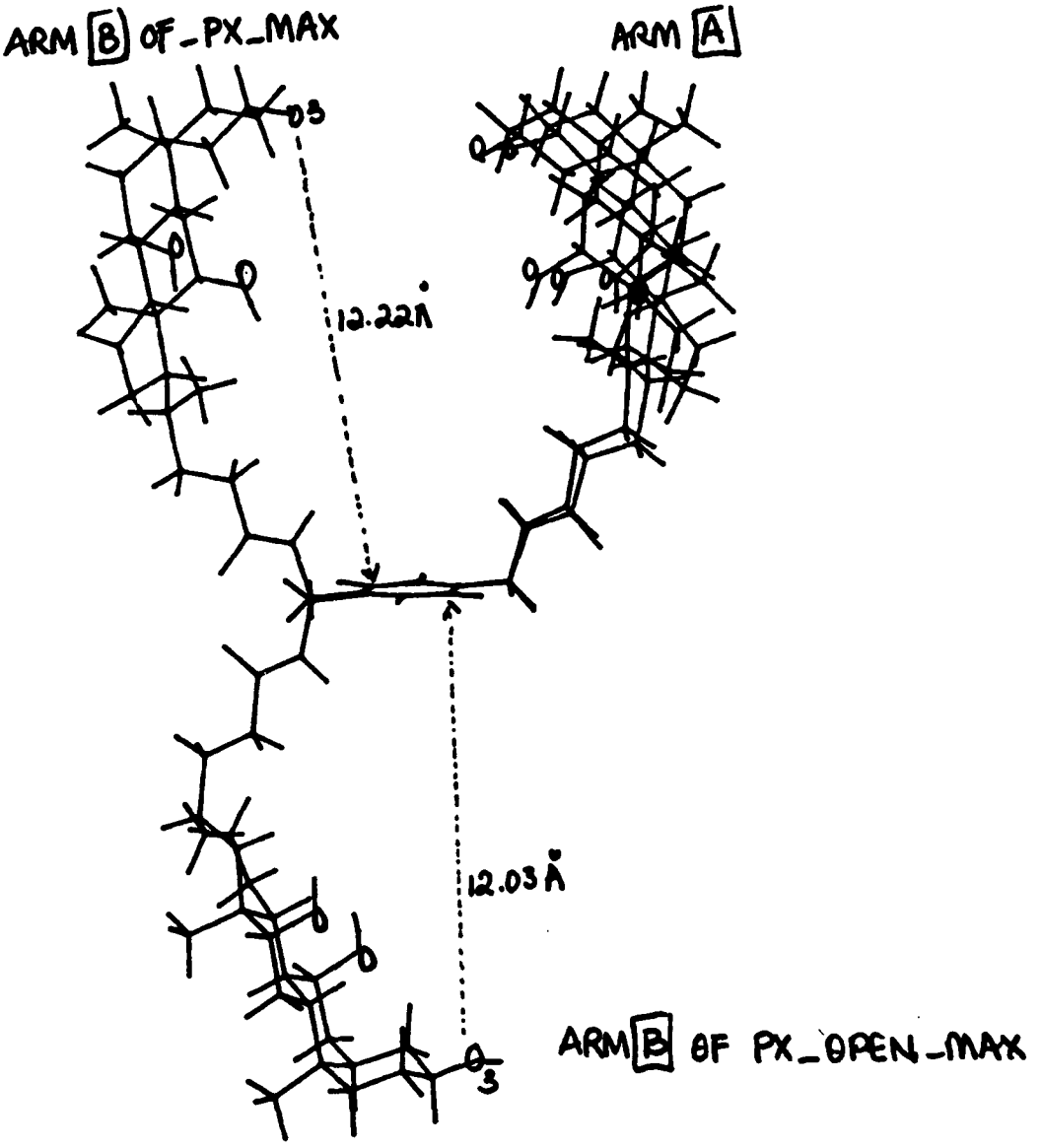


FIGURE 11.

CUBSUN_PX_~~MAX~~OC (BL) & CUBSUN_PX_OPEN_MAX